

JOURNAL OF MATHEMATICAL ANALYSIS AND APPLICATIONS **87**, 45–50 (1982)

A-Optimality for Regression Designs*

N. N. CHAN

*Department of Mathematics, The Chinese University of Hong Kong.
Shatin, N. T., Hong Kong**Submitted by K. Fan*

Use is made of a result of Horn (*Amer. J. Math.* **76** (1954), 620–630) on the existence of a symmetric matrix with prescribed diagonal elements and eigenvalues. A necessary and sufficient condition is then given for the existence of an A -optimal design for a regression experiment in the Dorogovcev (*Selected Transl. Math. Statist. Probab.* **10** (1971), 35–41) setting.

1. INTRODUCTION

Consider the linear regression model

$$y = X\beta + \varepsilon,$$

where y is an $m \times 1$ vector of observations, X is an $m \times n$ matrix to be called the design matrix, β is an $n \times 1$ vector of unknown parameters, and ε is an $m \times 1$ vector of random variables with mean the $m \times 1$ zero vector and known covariance matrix A . We assume that $m \geq n$ and denote the eigenvalues of A in ascending order of magnitude by

$$\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n \leq \dots \leq \lambda_m.$$

For later use denote the diagonal matrices with diagonal elements $\lambda_1, \dots, \lambda_i$ by A_i , $i = n$ and m .

For a given design matrix X of rank n , an unbiased estimate of the parameter β based on the observation y is the simple least squares estimate

$$(X'X)^{-1} X'y,$$

whose covariance matrix is given by

$$(X'X)^{-1} X'AX(X'X)^{-1}. \quad (1)$$

* This work was done while the author was a visiting scholar at Stanford University.

One of the design problems is to choose X from a given experimental region such that the trace of the matrix in (1) is minimal. This is a problem in the A -optimal designs of regression experiments and was considered by Dorogovcev [3] under the more general setting that the observations are the realization of stochastic processes. Earlier work on A -optimal designs was performed by Elfving [4] and Chernoff [2].

In this paper, the experimental region under consideration is taken to be the set H of all $m \times n$ real matrices of rank n whose i th column has a Euclidean norm not exceeding c_i , $i = 1, \dots, n$, where the c_i are given positive numbers. In Section 2, it is shown that for any matrix X in H the trace of the matrix in (1) has as a lower bound of

$$\left(\sum_{i=1}^n c_i^2 \right)^{-1} \left(\sum_{i=1}^n \lambda_i^{1/2} \right)^2.$$

In Section 3, a necessary and sufficient condition for the existence of an X in H to attain the lower bound is derived. For the case in which all the c_i are equal, a partial result was given in Chan and Wong [1]. Dorogovcev [3] obtained the lower bound for the special case $n = 2$ and $c_1 = c_2$.

It is worth noting that in the regression model if one considers the best linear unbiased estimate $(X'A^{-1}X)^{-1}X'A^{-1}y$ and its covariance matrix $(X'A^{-1}X)^{-1}$, by minimizing the trace of the latter for all X in H , the corresponding optimal design problem has a simple solution, as is given in Rao [10, p. 236]. On the other hand, if one wishes to minimize the determinant of $(X'A^{-1}X)^{-1}$, there is the so-called D -optimal design problem, of which comprehensive reviews can be found in St. John and Draper [11] and Kiefer and Galil [7].

2. AN INEQUALITY

For the regression model and the set H as given in Section 1, we note that in minimizing the trace of the matrix in (1) with respect to X in H , the matrix A in (1) can be replaced by the diagonal matrix A_m without loss of generality, in view of the existence of an orthogonal matrix P such that

$$A = P'A_mP$$

and the equality

$$(X'X)^{-1}X'A(X'X)^{-1} = (Y'Y)^{-1}Y'A_mY(Y'Y)^{-1},$$

where $Y = PX$, which is again in H . The following lemma of Fan [5] will be required in the proof of our main inequality.

LEMMA 1. Let B be a real $m \times n$ matrix whose n columns form an orthonormal set. Then

$$\text{tr } B'AB \geq \text{tr } A_n,$$

where tr represents the trace operation.

THEOREM 1. For any X in H ,

$$\text{tr} \{ (X'X)^{-1} X'AX(X'X)^{-1} \} \geq \left(\sum_{i=1}^n c_i^2 \right)^{-1} \left(\sum_{i=1}^n \lambda_i^2 \right)^2.$$

Proof. By the Cauchy-Schwarz inequality applied to the trace inner product $\text{tr} \{ X'Y \}$ between two real $m \times n$ matrices X and Y , we have

$$\text{tr} \{ X'X \} \times \text{tr} \{ (X'X)^{-1} X' A^{1/2} A^{1/2} X (X'X)^{-1} \} \geq \text{tr}^2 \{ X' A^{1/2} X (X'X)^{-1} \}. \quad (2)$$

But the trace on the right-hand side is

$$\text{tr} \{ (X'X)^{-1/2} X' A^{1/2} X (X'X)^{-1/2} \}, \quad (3)$$

which is not less than $\sum_{i=1}^n \lambda_i^{1/2}$ by Lemma 1 on noting that the n columns of the matrix $X(X'X)^{-1/2}$ are orthonormal. By the definition of the set H ,

$$\text{tr} \{ X'X \} \leq \sum_{i=1}^n c_i^2. \quad (4)$$

Hence the main inequality follows.

3. A-OPTIMAL DESIGNS

The main result of this work is to obtain a necessary and sufficient condition on $(\lambda_1, \dots, \lambda_m)$ and (c_1, \dots, c_n) for the existence of a matrix in H such that the lower bound in Theorem 1 is attained. For this we need the following lemmas.

LEMMA 2. Let D be an $n \times n$ real diagonal matrix with diagonal elements $d_1 \leq d_2 \leq \dots \leq d_n$, and a_1, \dots, a_n be n real numbers such that $a_1 \leq a_2 \leq \dots \leq a_n$ and

$$\sum_{i=1}^n a_i = \sum_{i=1}^n d_i.$$

Then there exists an $n \times n$ orthogonal matrix P such that the n diagonal elements of $P'DP$ are a_1, \dots, a_n if and only if

$$\sum_{i=1}^k a_i \geq \sum_{i=1}^k d_i, \quad k = 1, 2, \dots, n-1.$$

This lemma is a version of a result by Horn [6] and a proof is given by Mirsky [9]. See also Marshall and Olkin [8, p. 220].

LEMMA 3. Let D be as in Lemma 2 and B be an $n \times k$ matrix whose k columns form an orthonormal set. Arrange the eigenvalues of the $k \times k$ matrix $B'DB$ in ascending order $b_1 \leq b_2 \leq \dots \leq b_k$. Then $b_i \geq d_i$, $i = 1, \dots, k$.

This is the Poincaré separation theorem and can be found, for example, in Rao [10, p. 64].

THEOREM 2. Suppose that the positive numbers c_i , $i = 1, \dots, n$, are arranged in ascending order of magnitude and that the smallest eigenvalue λ_1 of the covariance matrix A is positive. Then there is an X in H such that

$$\text{tr}\{(X'X)^{-1}X'AX(X'X)^{-1}\} = \left(\sum_{i=1}^n c_i^2\right)^{-1} \left(\sum_{i=1}^n \lambda_i^{1/2}\right)^2$$

if and only if

$$\left(\sum_{i=1}^n c_i^2\right)^{-1} \sum_{i=1}^k c_i^2 \geq \left(\sum_{i=1}^n \lambda_i^{1/2}\right)^{-1} \sum_{i=1}^k \lambda_i^{1/2}, \quad k = 1, \dots, n-1.$$

Proof. Sufficiency. Consider the diagonal matrix $A_n^{1/2}$ whose diagonal elements are $\lambda_i^{1/2}$, $i = 1, \dots, n$. By Lemma 2 there exists an orthogonal matrix P of order n such that the i th diagonal element of $P'A_n^{1/2}P$ is bc_i^2 , $i = 1, \dots, n$, where

$$b = \left(\sum_{i=1}^n c_i^2\right)^{-1} \left(\sum_{i=1}^n \lambda_i^{1/2}\right).$$

Denote by X the $m \times n$ matrix

$$b^{-1/2} \begin{bmatrix} A_n^{1/4}P \\ O \end{bmatrix},$$

where O is an $(m-n) \times n$ submatrix of zeros. Note that X is of rank n as $\lambda_1 > 0$ and that the i th diagonal element of $X'X$ equals c_i^2 as we have

$$X'X = b^{-1}P'A_n^{1/2}P.$$

Hence X is a member of the set H . Moreover, for the diagonal matrix A_m of order m , we have

$$\begin{aligned} X' A_m X &= b^{-1} [P' A_n^{1/4} \quad O'] A_m \begin{bmatrix} A_n^{1/4} P \\ O \end{bmatrix} \\ &= b^{-1} P' A_n^{1/4} A_m A_n^{1/4} P \\ &= b^{-1} P' A_n^{3/2} P, \end{aligned}$$

and so

$$\begin{aligned} \text{tr}\{(X'X)^{-1} X' A_m X (X'X)^{-1}\} &= \text{tr}\{b(P' A_n^{-1/2} P) P' A_n^{3/2} P (P' A_n^{-1/2} P)\} \\ &= b \text{tr}\{P' A_n^{1/2} P\} \\ &= \left(\sum_{i=1}^n c_i^2\right)^{-1} \left(\sum_{i=1}^n \lambda_i^{1/2}\right)^2. \end{aligned}$$

The proof for sufficiency is completed by replacing A_m by A as remarked at the beginning of Section 2.

Necessity. Suppose that X , a member of H , is such that the inequality in Theorem 1 becomes an equality. Then the three inequalities in the proof of Theorem 1 reduce to equalities. First, note that the i th diagonal element of the matrix $X'X$ equals c_i^2 , $i = 1, \dots, n$, because it cannot exceed c_i^2 (as X is in H), and from (4),

$$\text{tr}\{X'X\} = \sum_{i=1}^n c_i^2.$$

By Lemma 2, it is then enough to show that $\lambda_1^{1/2}, \dots, \lambda_n^{1/2}$ are the eigenvalues of the $n \times n$ matrix $bX'X$. For this, note that the Cauchy-Schwarz inequality (2) becoming an equality implies that there is a nonzero real number d such that

$$X = dA^{1/2}X(X'X)^{-1}.$$

So we have

$$X'X = dX'A^{1/2}X(X'X)^{-1}.$$

The equality corresponding to (3) then implies that

$$\begin{aligned} \sum_{i=1}^n \lambda_i^{1/2} &= \text{tr}\{X'A^{1/2}X(X'X)^{-1}\} \\ &= d^{-1} \text{tr}\{X'X\} \\ &= d^{-1} \sum_{i=1}^n c_i^2. \end{aligned} \tag{5}$$

Therefore, $d = b^{-1}$, and so

$$bX'X = X'A^{1/2}X(X'X)^{-1}.$$

It remains to show that the $n \times n$ matrix

$$(X'X)^{-1/2} X'A^{1/2}X(X'X)^{-1/2} \quad (6)$$

has $\lambda_1^{1/2}, \dots, \lambda_n^{1/2}$ as its eigenvalues. In fact, by replacing A by A_m and using Lemma 3, we see that the i th smallest eigenvalue of the matrix in (6) is not less than $\lambda_i^{1/2}$, $i = 1, \dots, n$, and, in view of the first equality in (5), must be equal to $\lambda_i^{1/2}$, completing the proof.

ACKNOWLEDGMENT

I am thankful to Professor T. W. Anderson for helpful discussions and suggestions.

REFERENCES

1. N. N. CHAN AND C. S. WONG, Existence of an A -optimal model for a regression experiment, *J. Math. Anal. Appl.* **77** (1980), 403–415.
2. H. CHERNOFF, Locally optimum designs for estimating parameters, *Ann. Math. Statist.* **24** (1953), 586–602.
3. A. JA. DOROGOVCEV, Problems of optimal control of a regression experiment, *Selected Transl. Math. Statist. Probab.* **10** (1971), 35–41.
4. G. ELFVING, Optimum allocation in linear regression theory, *Ann. Math. Statist.* **23** (1952), 255–262.
5. K. FAN, On a theorem of Weyl concerning eigenvalues of linear transformations, I, *Proc. Nat. Acad. Sci. USA* **35** (1949), 652–655.
6. A. HORN, Doubly stochastic matrices and the diagonal of a rotation matrix, *Amer. J. Math.* **76** (1954), 620–630.
7. J. KIEFER AND Z. GALIL, Optimum weighing designs, in "Recent Developments in Statistical Inference and Data Analysis," pp. 183–190, North-Holland, Amsterdam, 1980.
8. A. W. MARSHALL AND I. OLKIN, "Inequalities: Theory of Majorization and Its Applications," Academic Press, New York, 1979.
9. L. MIRSKY, Matrices with prescribed characteristic roots and diagonal elements, *J. London Math. Soc.* **33** (1958), 14–21.
10. C. R. RAO, "Linear Statistical Inference and Its Applications," 2nd ed., Wiley, New York, 1973.
11. R. C. ST. JOHN AND N. R. DRAPER, D -optimality for regression designs: A review, *Technometrics* **17** (1975), 15–23.